

NJSD

中国（南京）软件开发者大会  
China (Nanjing) Software Developers Conference

2016

# 大数据风控技术应用探索

上海安硕信息技术股份有限公司 

(安硕信息, 300380)

# 公司简介-Amarsoft

## 企业规模

2001年成立，2014年创业板上市，人员超过1700人

当前总体控参投近15家科技公司

## 业务方向

专注于信贷资产管理和信用风险管理领域解决方案

业务咨询、系统开发、大数据服务、建模等

## 客户群体

200多家银行以及类银行大中型金融机构客户

数千家普惠金融机构客户

# 传统金融主要客户列表

## 国有银行

中国银行(中银富登)  
建设银行(建银村镇)  
交通银行

## 全国性 股份制银行

招商银行  
浦发银行  
民生银行  
中信银行  
兴业银行  
光大银行  
平安银行  
广发银行  
渤海银行  
浙商银行  
恒丰银行

## 区域性银行

### 直辖市：

北京银行 上海银行 天津银行 重庆银行

### 省级合并：

徽商银行 **江苏银行** 龙江银行 长安银行  
湖北银行 贵州银行

### 省会：

汉口银行 宁夏银行 富滇银行 河北银行  
**南京银行** 晋商银行 广州银行 西安银行  
兰州银行 盛京银行 贵阳银行 杭州银行  
福建海峡银行 内蒙古银行 乌鲁木齐商行  
哈尔滨银行 成都银行

### 计划单列市：

宁波银行 青岛银行 大连银行

### 央企控股银行

昆仑银行 华润银行

## 农村金融机构

**江苏农信** 湖北农信  
陕西农信 吉林农信  
北京农商 上海农商  
成都农商 天津农商  
广州农商 深圳农商  
南海农信 东莞农商  
长春农商

## 外资（合资）银行

花旗银行 东亚银行  
富邦华一银行

## 非银行金融机构

资产：人寿 信达 平安

信托：平安 华润 国民  
保险：安邦 友邦 太平洋  
证券：中信 平安



近200家国内外金融机构

11家全国性股份制商业银行

90余家国内大中型城市商业银行

50多家重要的非银行金融机构

# 新兴金融主要客户列表

## 互联网金融机构

互联网系：

腾讯财付通

微众银行

宜信

传统金融系：

证通股份

太平电商

安邦帮付宝

招联消费金融

佰仟金融

大地财保

平安陆金所

产业系：

国美电器

复星高科

中兴通讯

天虹百货

伊利集团

万达金融

永达金融

其他：

大数金融

汇金达

华融聚辉

厚本金融

平安租赁

## 普惠金融机构

大众小贷、张江小贷等近千家小贷公司和融资性担保公司

## 合作运营

齐商银行

潍坊农商行

易贷中国

江西省小贷协会

# 主要业务版图

## ■ 金融IT实施与咨询服务

信贷管理系统产品群  
额度、押品、核算系统  
风险审计管理产品群  
风险预警系统  
资产证券化系统  
.....

## ■ 安硕大数据服务

大数据技术服务  
大数据运营服务  
信用风险数据集市  
信用风险数据仓库  
信用风险规则仓库  
.....

金融IT  
大数据  
云服务

## ■ 安硕云服务平台

小贷云服务平台

大数据征信云服务平台

供应链金融云服务平台

网贷云服务平台

消费金融云服务平台

# 大数据风控技术应用探索



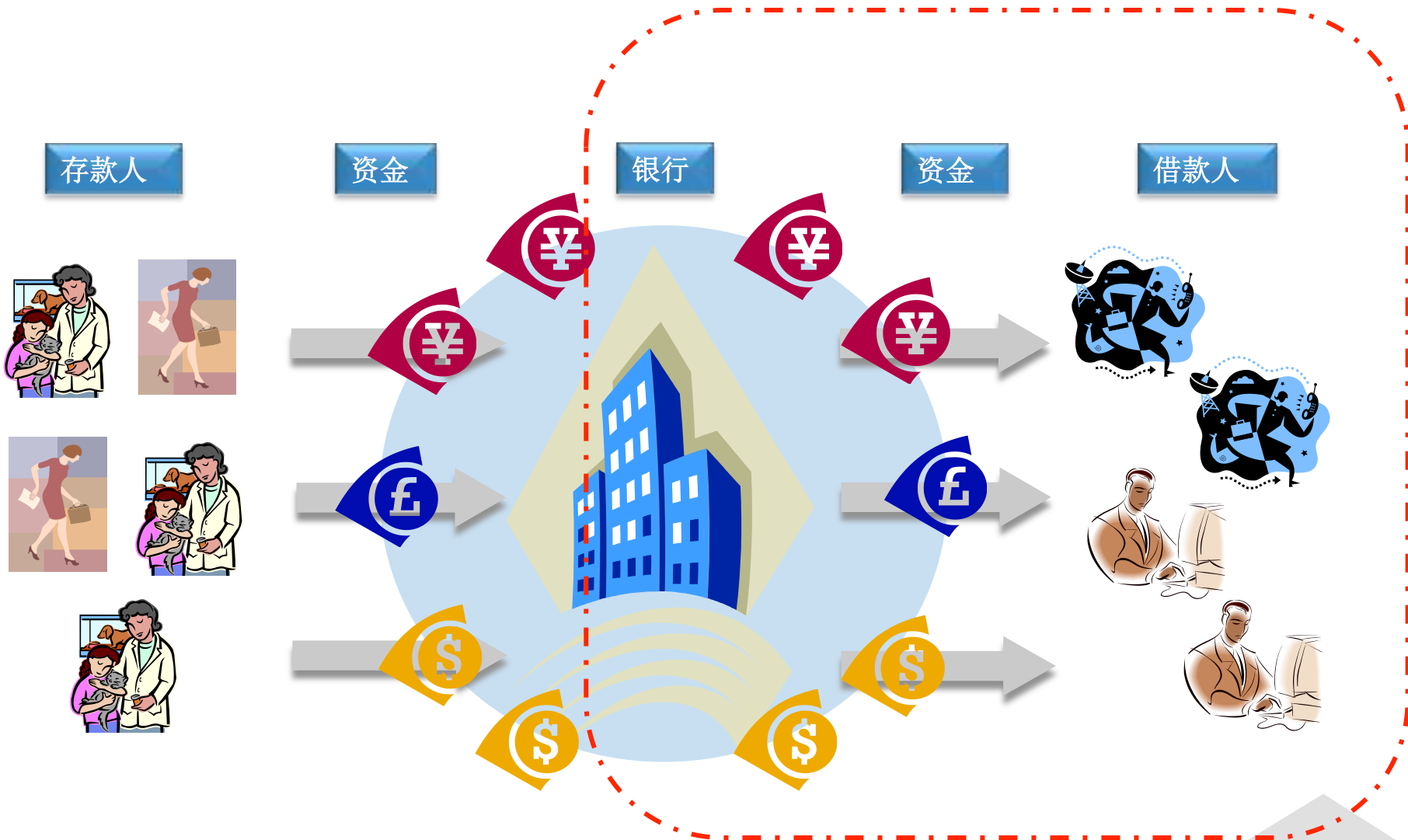


01

---

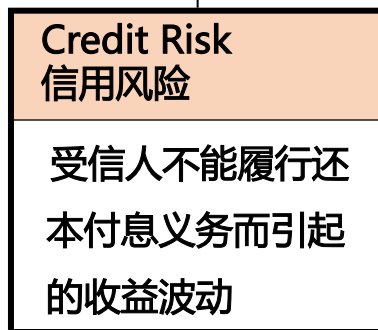
## 传统风控技术应用概述

# 什么是传统的商业银行运营模式？

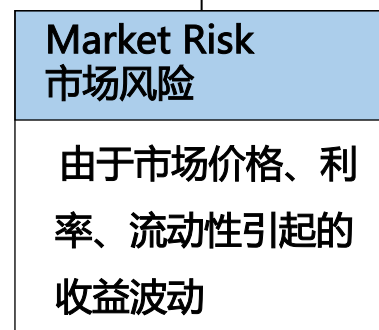




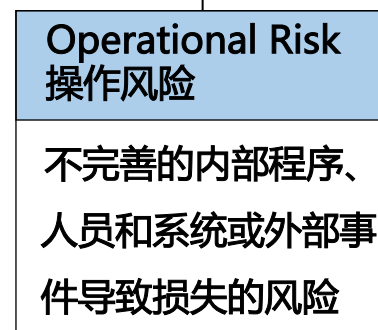
# 风险管理核心内容



- ❑ 意愿
- ❑ 偿付能力



- ❑ 利率风险
- ❑ 汇率风险
- ❑ 证券价格
- ❑ 流动性风险
- ❑ 隐含波动性



- ❑ 人员
- ❑ 流程
- ❑ 系统
- ❑ 外部事件

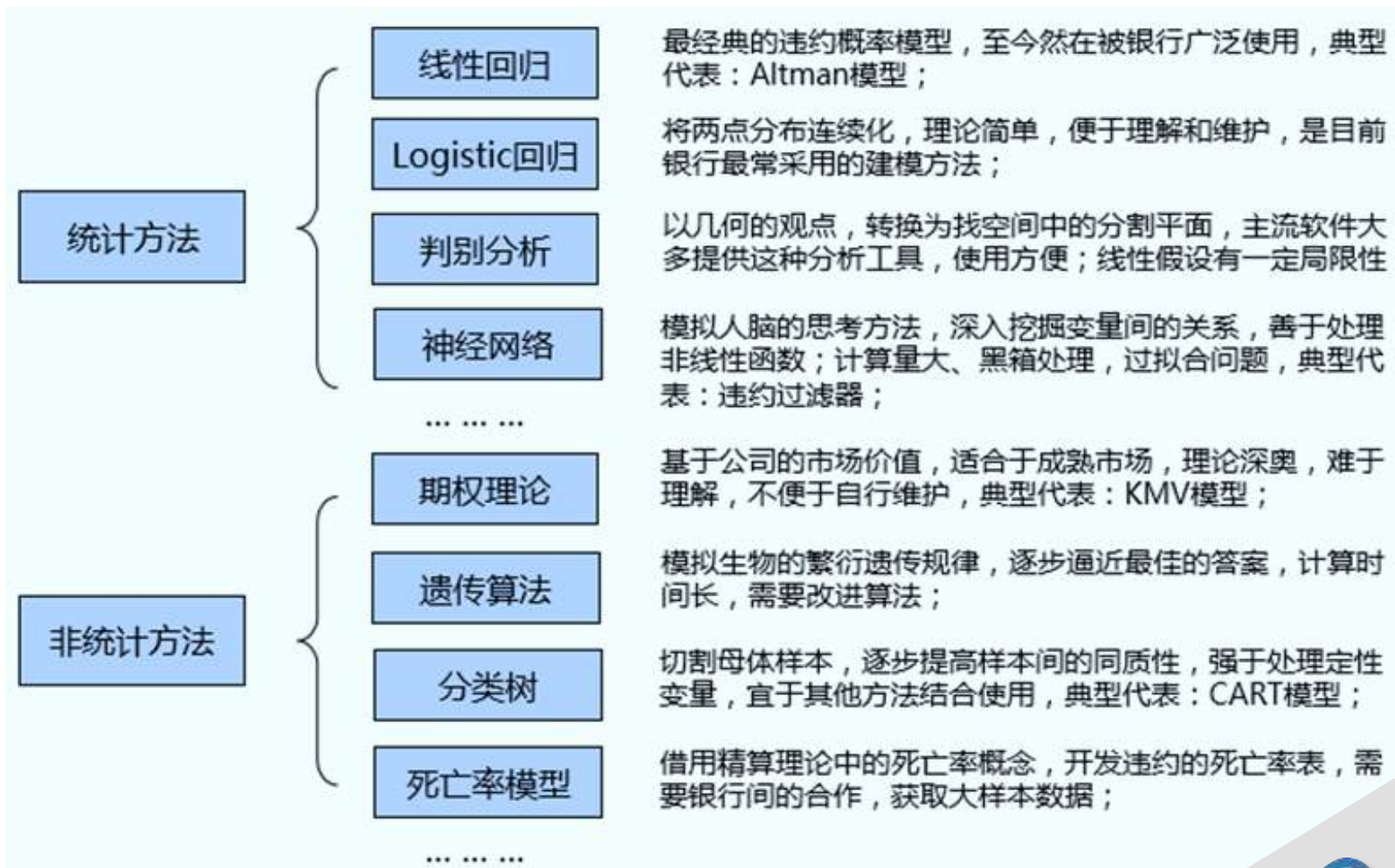
信用风险是银行最为复杂的风险种类，也是银行面临的最主要的风险

# 传统风控技术应用

## 信贷周期

## 模型设计

## 算法应用



# 传统风控技术利弊分析

## 传统风控的优势

- 理论基础完善，经过成熟市场的实证检验
- 模型结构简单透明，容易理解和沟通
- 基于银行内部的信贷数据，预测性较强

## 传统风控的缺点

- 单家机构的数据微乎其微
- 传统数据类型的有偏性：金融、财务、借贷行为类数据（全国只有30%的人群拥有可靠的借贷历史数据）
- 变量数目的有限性：10多个变量

# 新金融领域风控痛点

## 信息不对称导致欺诈频频发生

借款人

### 央行个人征信数据不足

3.5亿人有信用记录

### 问题平台频出



与去年同期相比，问题平台数量增长了**7.5倍**

为什么问题平台如此之多？



痛点3

# 新金融领域风控痛点

## 怎么将坏客户拒之门外

- **痛点** 坏客户挡在门槛外
- **苦处** 不清楚哪些客户是不好的，信息来源少，及时性不够
- **趋势** 建立名单库（灰名单库），并对名单进行分类，拒绝、警告、提示，对应的准入标准
- **信息需求** 信息多渠道，除日常能够爬下来的公检法税务环保海关工商等公开网站外，利用新的征信渠道如芝麻信用（集团交易数据、合作银行反馈数据、租车生活等信息、主流p2p数据）、以及其他途径进行匹配是否一致)

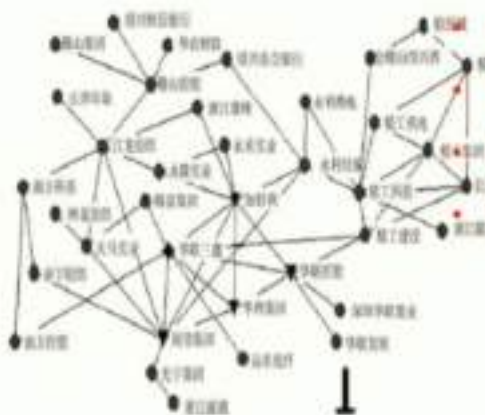


## 如何掌握上下游客户信息

- **痛点** 无法掌握那么细
- **苦处** 平时也不太关心，但从大数据分析看，上下游关联交易信息对客户本身影响最大最敏感
- **趋势** 依赖外部信息
- **信息需求** 引入外部数据



## 如何识别担保链



- **痛点** 说不清楚
- **苦处** 担保链实在复杂
- **趋势** 图谱化
- **信息需求** 将客户本身、法人代表、实际控制人等按一定图谱展示，并汇总信息

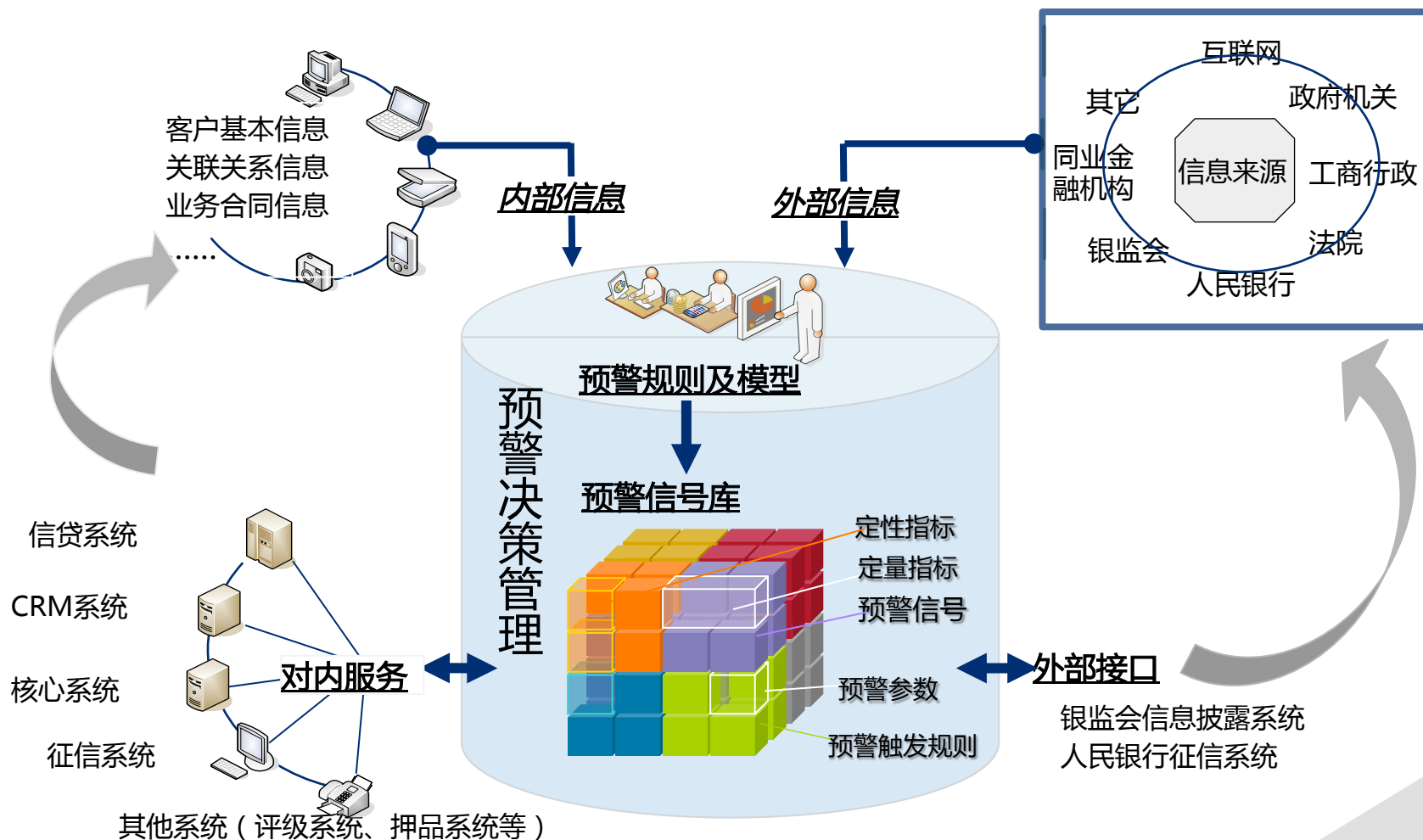
02

---

大数据风控技术应用探索

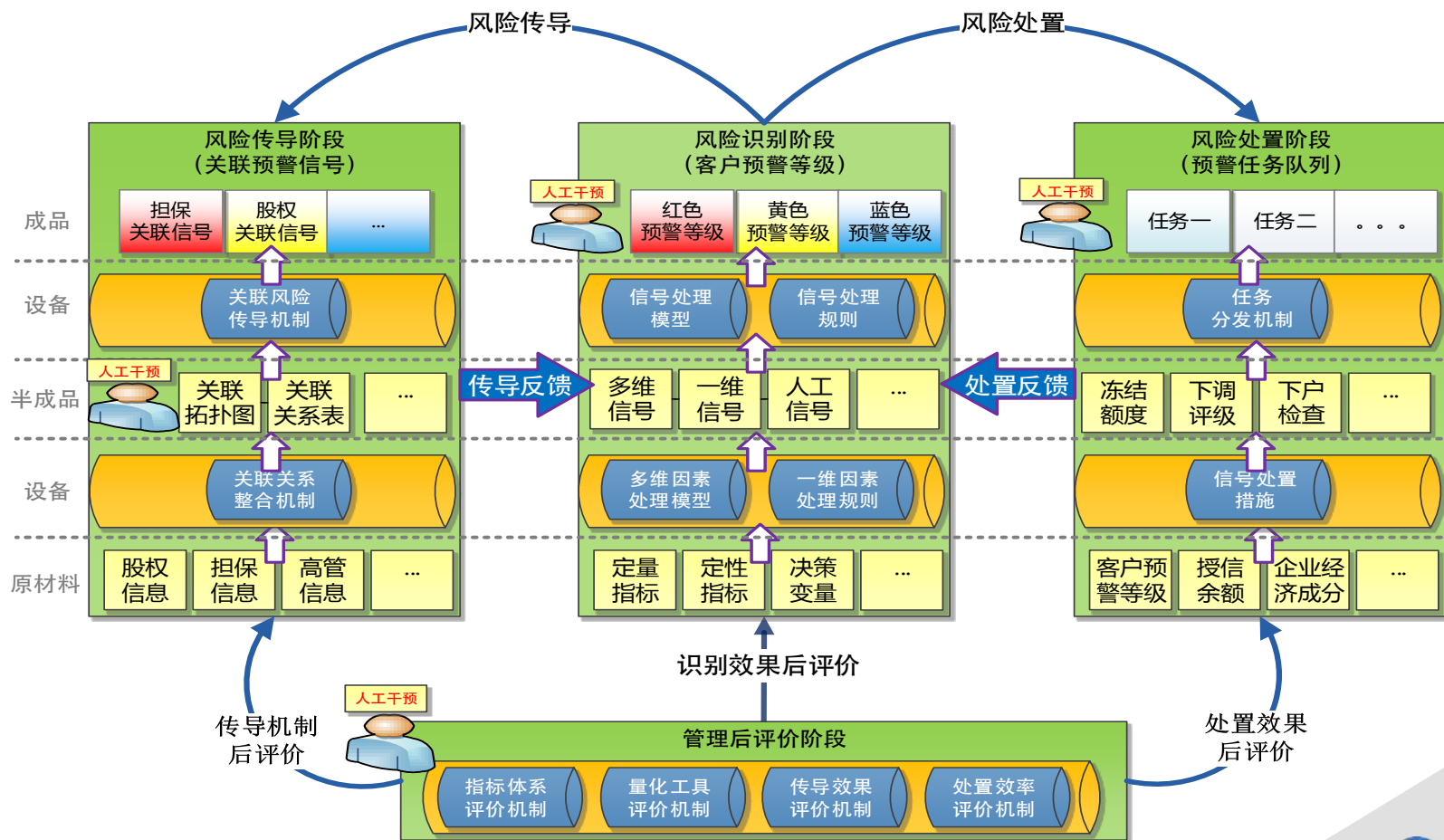
# 大数据风控技术要有前瞻性

## ➤ 内外部风险信息源接入管理



# 大数据风控技术要有先进性

- 先进性主要体现在**数据规划整合**、**数据分析挖掘**、**智能决策管理**、**决策结果应用**、**决策效果后评价**五个方面。





# 智能化决策管理机制

## 决策管理机制

根据客户虚警和漏警的情况重检策略，开展策略后评价工作，总结策略优化的方向和内容，对策略适用性进行持续动态维护。

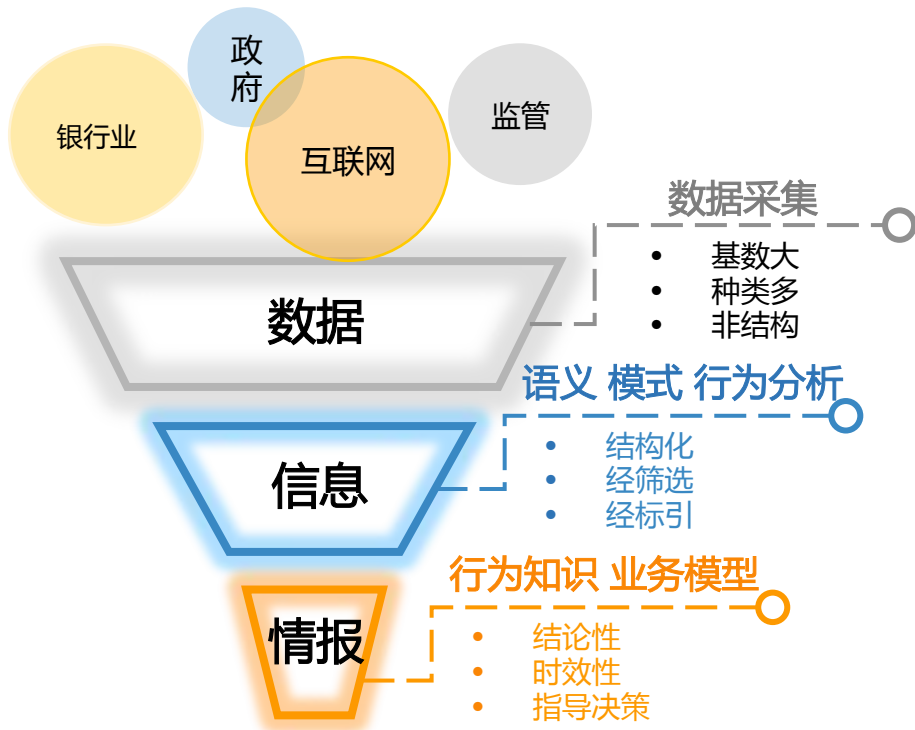
定义行内及行外风险预警信息源，根据数据质量分析报告，确定风险预警目标数据需求，提出数据治理建议

预警对象的定义需体现预警全面性的要求，根据风险因素定义预警主题，完善的预警主题有助于对信号进行清晰、明确的归集、识别和管理

在客户层面进行信号归集和排重，综合评定客户风险等，结合风险偏好设计推送策略和跟踪处置建议

定义有重大影响和控制力的关联关系；合理划分关系类型；确定关联关系管理目标数据需求；设计关联风险传导机制

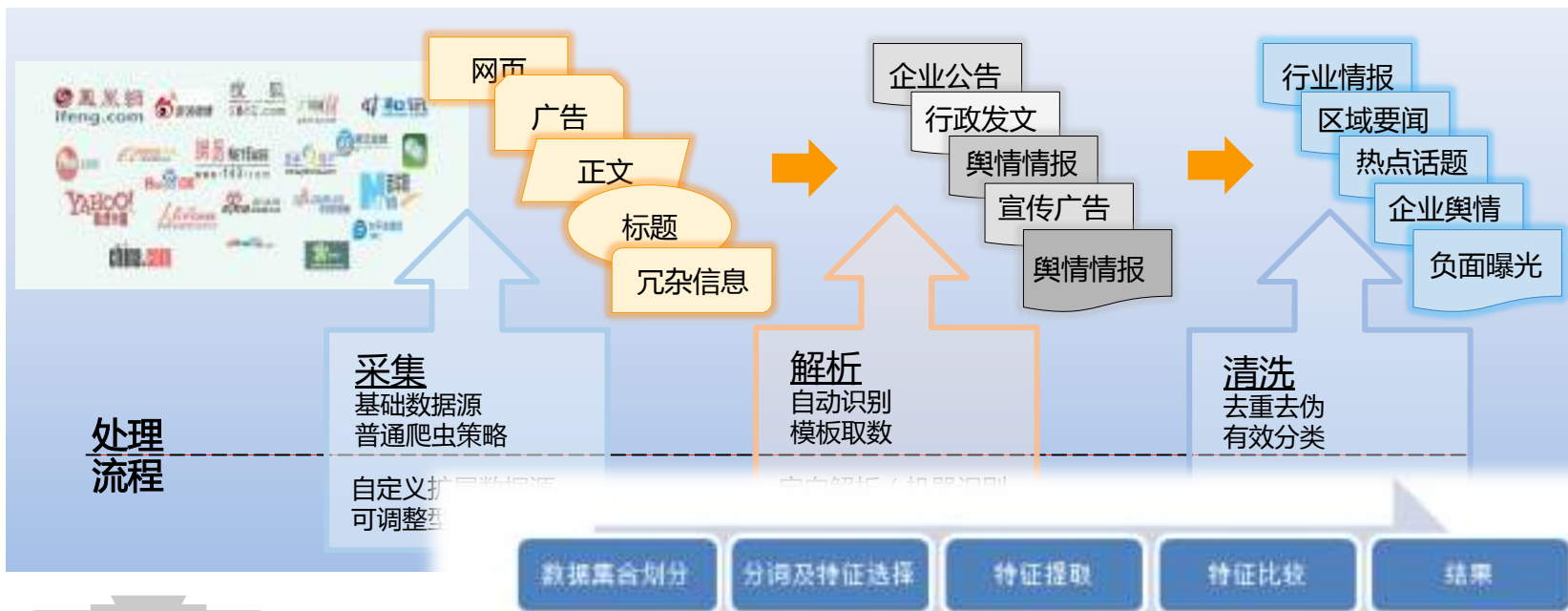
# 核心问题——大数据处理诸多挑战



- 数据特性：数据源体现多源、多态、多构，体量大、碎片化的特点
- 数据处理：采集、存储、解析、去重、降噪、结构化、整合难度大
- 数据组织：需要建立各群体（行业、项目、集团、企业、个人）广度及深度关联

■ IT架构保障：传统架构VS互联网云架构OR混合架构

# 非结构化数据处理过程示例



## 舆情专家审核

对象匹配判读、策略匹配判读



命中企业：企业A

文章标题：\*\*\*\*，\*\*\*\*行业面临重大考验

聚焦情感倾向：负面

一级主题：经营管理

新闻来源：\*\*

预警等级：高

二级主题：经

分词算法	特征选择方法	哈希函数	特征提取
iKA	词频	Times33	SimHash
HPL	权重	TF/IDF	MinHash
极易分词		TestRank	投影Hash
	词频+权重		SimHash+向量余弦
	shingling		n-gram
	特征码(句号前后)		LSA/PALSA

# 智能结构化示例

## 未结构化数据

### 宁夏宁电硅材料有限公司与安徽蓝德集团股份有限公司、中铝宁夏能源集团有限公司买卖合同纠纷二审民事裁定书

提交时间: 2014-05-26 打印本文 查看全文

#### 安徽省高级人民法院 民事裁定书

(2014)皖民二终字第00237号

上诉人(原审被告): 宁夏宁电硅材料有限公司。

法定代表人: 王海勃, 该公司总经理。

被上诉人(原审原告): 安徽蓝德集团股份有限公司。

法定代表人: 李正祥, 该公司董事长。

原审被告: 中铝宁夏能源集团有限公司。

上诉人宁夏宁电硅材料有限公司因与被上诉人安徽蓝德集团股份有限公司、原审被告中铝宁夏能源集团有限公司买卖合同纠纷一案, 不服安徽省滁州市中级人民法院于2013年12月19日作出的(2013)滁民二初字第00147-2号民事裁定, 向本院提出上诉称: 2010年4月23日、5月25日, 上诉人与被上诉人分别签订两份供货合同, 明确约定“因合同引起或与合同有关的任何争议, 双方协商不成的, 由需方(即上诉人)所在地人民法院管辖”。在合同履行过程中, 被上诉人向上诉人发出对账函, 督促上诉人付款, 虽然被上诉人在对账函中写有若上诉人在规定的时间内不付款, 被上诉人可在其所在地法院诉讼, 但是上诉人仅对应收账款金额作出确认, 并未认可变更管辖, 请求撤销原裁定, 依法将本案移送宁夏回族自治区银川市西夏区人民法院审理。

本院经审理查明, 安徽蓝德集团股份有限公司因合同货款向宁夏宁电硅材料有限公司发出账务核对函, 函中写有“如不付款, 由我公司所在地人们法院裁决”, 函中“人们”二字被划掉, 宁夏宁电硅材料有限公司在账务核对函上加盖了公司公章, 并手写“数据无误金学刚2013年8月5日”文字, 根据函中内容可认定双方当事人对管辖法院做了约定, 即便如上诉人所称此前已有管辖法院约定, 那么也可认为此账务核对函对原有争议解决方式做了重新约定。依据法律规定, 合同纠纷的当事人可以书面协议选择原告住所地、合同履行地或与争议有实际联系的地点的人民法院管辖, 现安徽蓝德集团股份有限公司根据双方最后约定向其公司所在地即原告住所地人民法院安徽省滁州市中级人民法院提起诉讼, 符合法律规定, 上诉人宁夏宁电硅材料有限公司上诉理由不能成立, 安徽省滁州市中级人民法院依法对本案具有管辖权。依据《中华人民共和国民事诉讼法》第三十四条、第一百七十条第一款第(一)项、第一百七十一条、第一百七十五条的规定, 裁定如下:

驳回上诉, 维持原裁定。

本裁定为终审裁定。

审 判 长 王永安  
代理审判员 李家群  
代理审判员 岳 磊  
二〇一四年五月十四日

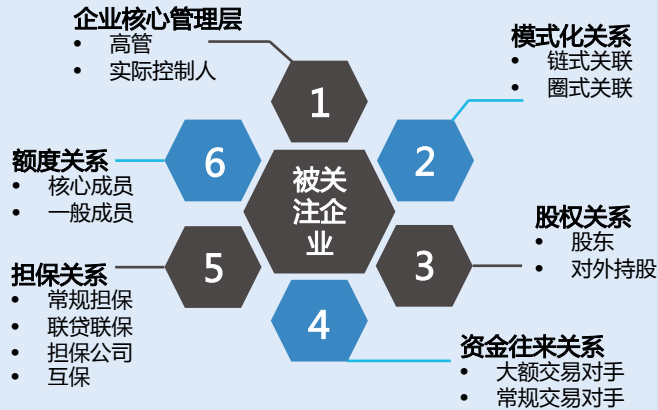
## 结构化数据

序号	2014101700003648
被执行人	安徽蓝德集团股份有限公司
案号	(2014)皖民二终字第00237号
文书类型	民事案件
公告类型	裁判文书
案件标题	宁夏宁电硅材料有限公司与安徽蓝德集团
审判长	刘伟芳
审判员	李建国
书记员	冯丹
立案日期	2013年12月19日
发布法院	安徽省高级人民法院
省份	安徽
城市	合肥市
公告日期	2014/05/26
公告地址	<a href="http://www.court.gov.cn/zgcpwsw/a">http://www.court.gov.cn/zgcpwsw/a</a>
数据来源	中国裁判文书网

示例

# 不同主体广度/深度关联挖掘

## 企业主体



客户关联关系管理拓扑图



## 个人主体



# IT架构保障挑战

## 优势

- 1、分布式水平扩展架构，容量扩展限制少
- 2、框架开源程度高，技术相对安全可控
- 3、软硬件产品总拥有成本（TCO）较低

## 互联网云架构

非核心账务类系统

## 痛点

- 1、商业化能力依赖于企业定制，对企业的开发能力要求高
- 2、强一致性的保障依赖场景的定制设计，并需要业务策略的支撑，对应用层要求较高
- 3、基于互联网技术生态解决方案的产品链缺口较大，商业化解决方案体系正持续完善中

核心账务类系统

## 传统架构

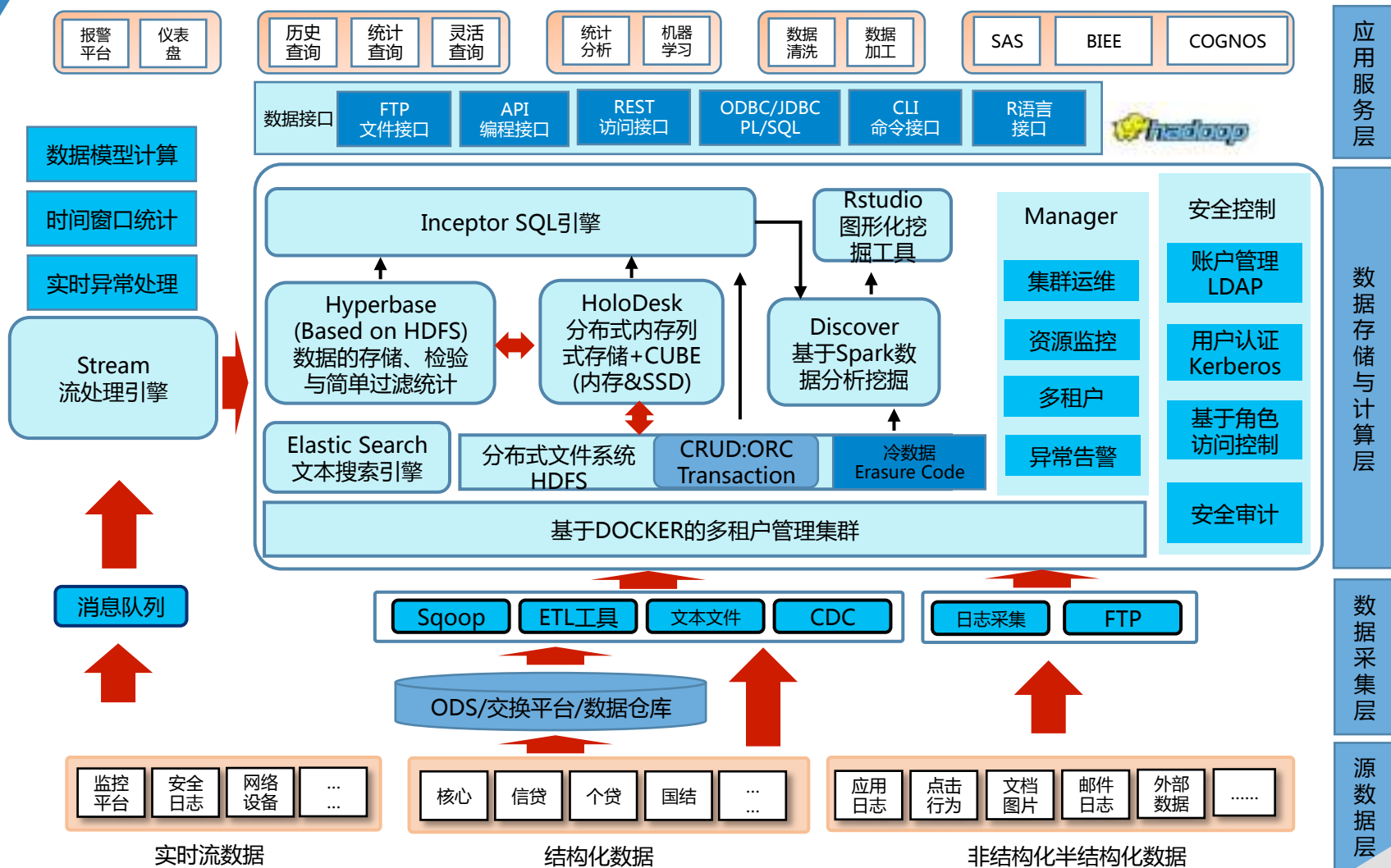
## 优势

- 1、产品商业化程度高，产品链、解决方案体系完善
- 2、满足金融行业的交易强一致性要求
- 3、安全性、合规性、可行性经过时间验证，充分满足监管要求

## 痛点

- 1、扩展性较差，依赖于单个硬件的纵向扩展
- 2、产品、技术相对封闭，功能扩展不灵活
- 3、软硬件产品总拥有成本（TCO）相对较高

# 基础技术平台架构



# 基础技术平台与应用集成混合架构

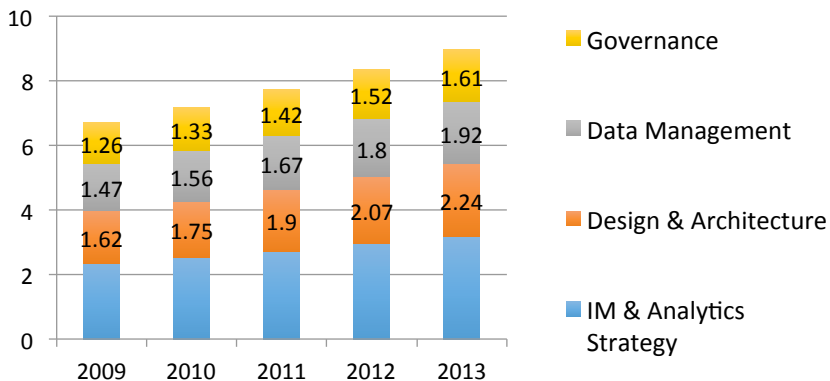




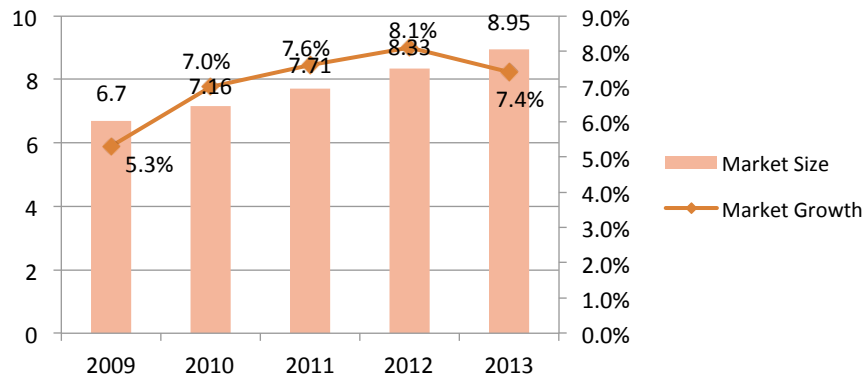
# 信息管理重要性&金融行业投入

➤ Gartner 的市场调研揭示，以目前趋势2017年很可能是企业信息危机爆发的时期，有意识的企业已经开始提升其信息管理的的能力规避信息风险

## ■ 数据管理是企业信息管理的基石



## ■ 数据管理和分析应用逐步成为IT建设的核心内容



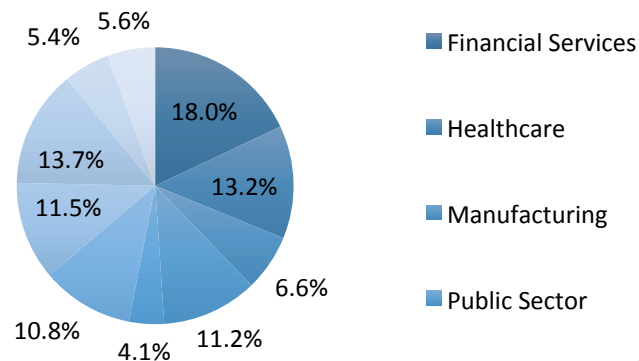
## ■ 大数据将进一步推动数据治理、数据管理的进程

Year	Enterprise Software Spending for Specified Sub-Markets	Forecast: Social Media Revenue, Worldwide, 2011-2016	Big Data IT Services Spending	Total
2011	2,565	76	24,407	27,047
2012	2,918	1,364	23,476	27,778
2013	3,516	1,812	28,578	33,906
2014	4,240	2,827	37,404	44,472
2015	5,207	3,615	36,199	45,010
2016	6,481	4,411	43,713	54,605

Note: Accuracy is to the nearest \$1 million and is derived from percentage-based algorithms. Research data does not provide total accuracy to the nearest \$1 million. Some rounding errors apply.

Source: Gartner (October 2013)

## ■ 金融行业是信息管理投入最高的行业



# 谢谢！

安硕信息 | [www.amarsoft.com](http://www.amarsoft.com)



## 上海

杨浦区国泰路11号复旦大学科技园大厦A座 23 楼  
邮编：200433  
电话：8621-35885888

## 重庆

重庆市沙坪坝区西永微电园西园南街5号R26-2号研发楼8层  
邮编：401332  
电话：8623-86259860

## 北京

东城区安定门东大街28号1号楼（雍和大厦）C座 801  
邮编：100007  
电话：8610-66218398

## 深圳

深圳市福田区车公庙耀华创建大厦 1904-1905  
邮编：518042  
电话：86755-23957377

## 苏州

苏州市苏州科技城科灵路 78 号  
邮编：215000  
电话：86512-68075325

## 厦门

厦门市思明区软件园观日路 32 号 303 单元 02 区  
邮编：361008  
电话：86592-2930996